

Orthogonal Riboswitches for Tuneable Coexpression in Bacteria**

Neil Dixon, Christopher J. Robinson, Torsten Geerlings, John N. Duncan,
Sheona P. Drummond, and Jason Micklefield*

Traditional small-molecule-inducible expression systems, such as those derived from the lactose (*lac*), arabinose (*ara*) and tetracycline (*tet*) operons, have been extensively utilized for the production of recombinant proteins,^[1] gene function analysis,^[2] target validation, and drug discovery.^[3] Inducible expression systems have also become indispensable tools in synthetic biology and have been used to control genetic circuits and networks,^[4] as well as metabolic pathways for the production of pharmaceuticals, biofuels, and fine chemicals.^[5] Despite these and many other notable successes, the commonly used regulatory systems do have limitations. Some of these systems exhibit “all-or-none” expression profiles, where a subpopulation of cells is fully induced, while others remain uninduced.^[6] Furthermore, some systems, such as *lac* and *ara*, exhibit “cross-talk”, which limits their use in applications where it is desirable to control the simultaneous and differential expression of multiple genes.^[7] In addition to synthetic biology, the controlled expression of multiple genes has become increasingly important for coexpression of multi-protein complexes,^[8] and to investigate protein–protein interactions.^[9] Given the incompatibility of some inducible promoters, and the difficulties that are associated with predicting the relative strengths of constitutive promoters, new approaches are urgently required to enable balanced expression of multiple proteins.^[10]

As an alternative to traditional protein-based expression systems, we recently reported orthogonal riboswitches^[11] that are produced by re-engineering a natural adenine-sensing riboswitch from the *add* gene of *Vibrio vulnificus*.^[12] The *add* A riboswitch is a “translational ON” switch that activates gene expression in the presence of its ligand inducer (Figure 1) through the disruption of a repressor stem, which otherwise sequesters the ribosome binding site (RBS). Through site-directed mutagenesis and a comprehensive screening strategy, we successfully created new RNA regulatory elements, which no longer respond to the natural metabolites, but which can instead be controlled by the addition of specific synthetic molecules^[11] (Figure 1). In this

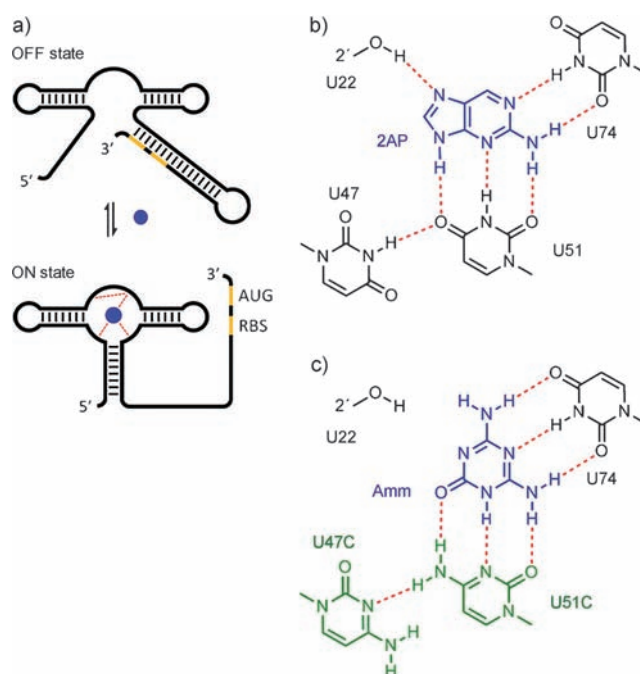


Figure 1. a) Schematic of the translational mechanism of gene regulation by the *add* A riboswitch. In the absence of inducer the OFF state is favored, in which the RBS and start codon (AUG) are sequestered in a repressor stem-loop. Binding of the small-molecule inducer (blue circle) stabilizes the ON state, which releases the RBS sequence to initiate translation of the downstream mRNA. b) Hydrogen-bond contacts between the *add* A riboswitch and 2AP. c) Hydrogen-bond contacts between the M6-riboswitch and Amm. The nucleotides mutated in the ligand-binding site of M6 are depicted in green. Hydrogen-bonding models are proposed on the basis of published crystal structures of related riboswitch–ligand pairings.^[11,14]

study, we demonstrate how mutually orthogonal riboswitches can be coupled together and deployed in the same bacterial cell to independently control the coexpression of multiple genes in a dose-dependent response to distinct synthetic small molecules. The multicomponent constructs that we have developed allow dynamic expression landscapes to be accessed and optimal protein coexpression stoichiometries to be established, which may prove useful for studying gene networks, for applications of protein coexpression, and for other applications in synthetic biology.

In the first instance, a Dual Promoter system was constructed, in which the orthogonal M6 riboswitch, selected previously,^[11] controlling translation of the red fluorescent protein DsRed in response to the inducer ammeline (Amm). The bacteriophage λ *t_{RI}* transcription terminator^[13] was inserted downstream of the *DsRed* gene, followed by

[*] Dr. N. Dixon,^[†] Dr. C. J. Robinson,^[†] Dr. T. Geerlings,
Dr. J. N. Duncan, Dr. S. P. Drummond, Prof. J. Micklefield
School of Chemistry & Manchester Interdisciplinary Biocentre
The University of Manchester
131 Princess Street, Manchester M1 7DN (UK)
E-mail: jason.micklefield@manchester.ac.uk

[†] These authors contributed equally to this work.

[**] This work was supported by the BBSRC (Grants BB/D005612/1 & BB/I012648/1).



Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/ange.201109106>.

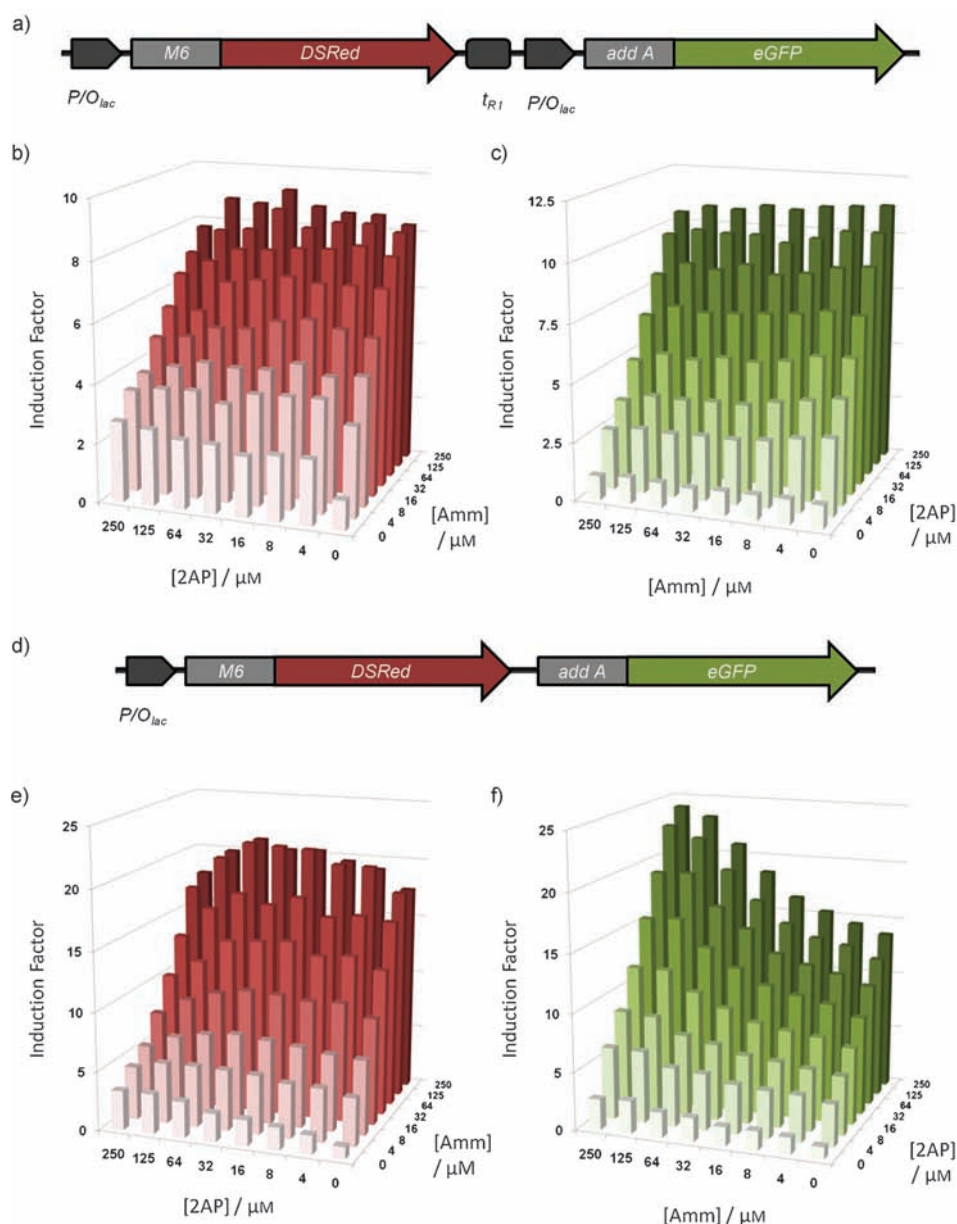


Figure 2. Riboswitch gene expression constructs and bivariate-dose gene expression matrices. a) The Dual Promoter construct comprised of a *lac* promoter/operator (P/O_{lac}), the Amm-responsive M6 riboswitch in front of the *DsRed* fluorescent protein gene, the phage λ tR1 transcriptional terminator, a second *lac* promoter/operator and the 2AP-responsive *add A* riboswitch in front of the *eGFP* fluorescent protein gene. b) Gene expression matrix of protein induction factors for Dual Promoter *DsRed* expression. c) Gene expression matrix for Dual Promoter *eGFP* expression. d) The Operon construct comprised of a single *lac* promoter/operator, the M6 riboswitch in front of the *DsRed* gene, followed directly by the *add A* riboswitch in front of the *eGFP* gene. e) Gene expression matrix for Operon *DsRed* expression. f) Gene expression matrix for Operon *eGFP* expression.

a second *lac* promoter and the 2-aminopurine (2AP)-responsive parental *add A* riboswitch,^[11,12] which controls translation of the green fluorescent protein *eGFP* (Figure 2a). *Escherichia coli* cells were transformed with this Dual Promoter construct, and expression of the *DsRed* and *eGFP* genes was monitored by fluorescence.

To explore the capacity of the Dual Promoter system for controlling the differential expression of *DsRed* and *eGFP*, a bivariate-dose gene-expression analysis was performed. In

this analysis, the induction factors (IFs), the ratio of ligand-induced protein expression over the basal level, were determined for a wide range of inducer concentrations (Figure 2). After co-administration of Amm and 2AP, it can be seen that the induction of *DsRed* expression increases in a dose-dependent manner in response to the Amm concentration, and reaches an induction factor of (9.3 ± 1.4) at $250 \mu\text{M}$ (errors given are SD). In contrast, the presence of 2AP over a range of concentrations has essentially no effect on *DsRed* expression (Figure 2b). Similarly, dose-dependent regulation of *eGFP* occurs in response to 2AP, which reaches an induction factor of (11.3 ± 0.5) at $250 \mu\text{M}$, with changes in Amm concentration having no effect upon *eGFP* levels (Figure 2c). These results show that the two riboswitches, despite being similar in sequence and structure, have excellent *in vivo* orthogonality when present within the same biological construct, and can be used to independently control the expression of the two genes simultaneously.

In a second series of experiments, a synthetic Operon system was constructed, in which one *lac* promoter drives the transcription of a single bicistronic mRNA, which comprises the M6 riboswitch for regulating translation of the *DsRed* gene, and an intergenic *add A* riboswitch that controls translation of the downstream *eGFP* gene (Figure 2d). A bivariate-dose gene-expression analysis demonstrates that differential control of *DsRed* with

an induction factor of (21.1 ± 2.9) can be achieved at a concentration of $250 \mu\text{M}$ of Amm (Figure 2e). However, analysis of *eGFP* expression showed interesting and unanticipated results. In this case, *eGFP* output is clearly dependent on both inducers. An apparent synergistic effect of the two inducers operates in the gene expression matrix, across the range of Amm and 2AP concentrations (Figure 2f). The *eGFP* expression landscape peaks at an induction factor of (24.1 ± 0.6) in the presence of $250 \mu\text{M}$ concentrations of both

inducers, whereas 250 μM of 2AP alone yields an induction factor of just (13.8 ± 0.4) . Our earlier in vitro studies clearly show that there is no binding interaction between Amm and the *add A* riboswitch, or between 2AP and M6.^[11] Similarly no “cross-talk” is detected with the Dual Promoter system described above. This suggests that *eGFP* expression within the Operon system is not only dependent on the intergenic *add A* riboswitch, but also on the upstream M6 riboswitch.

This effect was further investigated in a dose-response analysis (Figure 3). Titration of the M6-selective inducer Amm (0–1000 μM) against the Dual Promoter in either the

Figure 3. Dose-response curves for riboswitch-induced fluorescent protein expression. a) Dual Promoter DsRed expression. b) Dual Promoter eGFP expression. c) Operon DsRed expression. d) Operon eGFP expression. Scatter plot data represents the mean of three repeats. Amm concentration range (0–1000 μM) in the absence (empty triangles) or presence (filled triangles) of 2AP (250 μM); 2AP concentration range (0–1000 μM) in the absence (empty circles) or presence (filled circles) of Amm (250 μM). Curve fitting is described in the Supporting Information and fitting parameters are displayed in Table S1. F_{norm} are fluorescence units normalized for optical density at 620 nm.

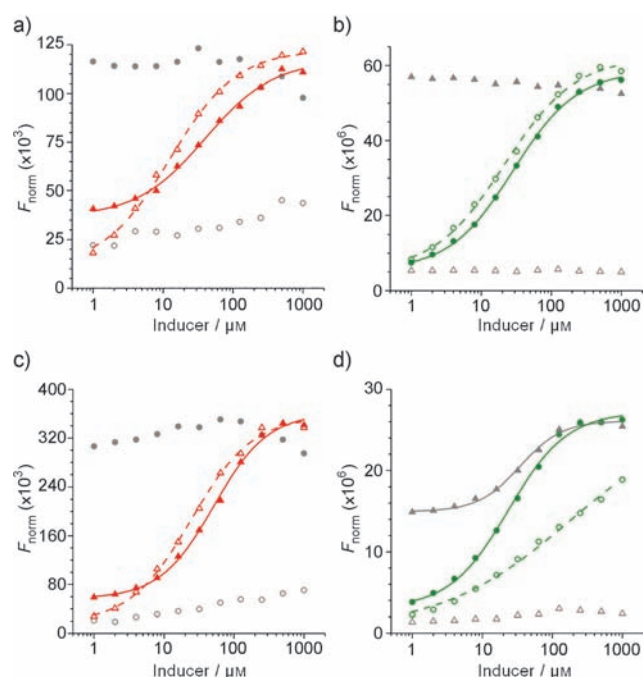


Figure 3. Dose-response curves for riboswitch-induced fluorescent protein expression. a) Dual Promoter DsRed expression. b) Dual Promoter eGFP expression. c) Operon DsRed expression. d) Operon eGFP expression. Scatter plot data represents the mean of three repeats. Amm concentration range (0–1000 μM) in the absence (empty triangles) or presence (filled triangles) of 2AP (250 μM); 2AP concentration range (0–1000 μM) in the absence (empty circles) or presence (filled circles) of Amm (250 μM). Curve fitting is described in the Supporting Information and fitting parameters are displayed in Table S1. F_{norm} are fluorescence units normalized for optical density at 620 nm.

absence or presence of 2AP (0 or 250 μM), gives dose-dependent protein expression curves for DsRed that are very similar, with maximum normalized red fluorescence levels of approximately 1.2×10^5 normalized fluorescence units (Figure 3a). Dose-dependent control of eGFP expression also occurs upon addition of the *add A* riboswitch inducer 2AP (Figure 3b). In this case, eGFP expression reaches a maximum level of approximately 6.0×10^7 normalized fluorescence units and is also largely unaffected by the presence of the M6 inducer Amm. For the Operon construct, titration of Amm (0–1000 μM) in the presence or absence of 2AP (0 or 250 μM) produces essentially identical dose-response curves for DsRed expression, where a maximum output of approxi-

mately 3.6×10^5 normalized fluorescence units is obtained (Figure 3c). In contrast, the synergistic effect for eGFP expression described above is clearly apparent, whereupon 2AP alone at a concentration of 1 mM produces an output of 1.89×10^7 normalized fluorescence units, whereas the same concentration of 2AP in the presence of Amm (250 μM) produces a maximal output of approximately 27.2×10^6 (Figure 3d). Furthermore, a dose-response relationship between eGFP expression and Amm concentration is clearly demonstrated when Amm is titrated in the presence of 250 μM of 2AP (Figure 3d). This confirms that both riboswitches exert an effect upon expression of the downstream *eGFP* gene within the Operon construct.

A number of reasons could account for the results obtained. The protection of full-length transcripts from nuclease degradation by actively translating ribosomes,^[15] which are recruited at both RBSs, would serve to increase mRNA stability in the presence of both inducers. In many regards, the effect upon eGFP expression mirrors the polar effect described for downstream genes in naturally occurring bacterial operons when translation of an upstream cistron is compromised.^[16] Polar effects have been shown to be associated with the prokaryotic termination factor Rho.^[17] In fact, riboswitches that control gene expression through a translational mechanism (like the *add A* riboswitch), might also rely on Rho-dependent termination of transcription.^[12b,18] Considering the expression of eGFP within the Operon system, when the upstream mRNA is not being actively translated (that is, in the absence of the M6 inducer Amm), the RNA polymerase must transcribe approximately 1000 nucleotides of mRNA before the eGFP RBS is reached, which provides ample opportunity for the recruitment of Rho and termination of transcription. Analysis of the mRNA levels for the single-component *add A*-eGFP construct by qPCR (Figure S1 in the Supporting Information) indeed showed that, upon addition of 2AP (250 μM), the mRNA levels doubled relative to transcript levels in the absence of 2AP (Figure S2 in the Supporting Information), which indicates that active translation is necessary to maintain high levels of transcript within the cells.

To further analyze the dose-dependent control of gene expression, the data for both constructs were fitted to a four-parameter logistic function. For the Dual Promoter construct (Table S1a in the Supporting Information), half-maximal induction (EC_{50}) was obtained with inducer concentrations in the 12–40 μM range. The dynamic range of inducer response (EC_{10} to EC_{90}) and Hill coefficients are consistent with independent coexpression of the two riboswitch-controlled genes. For the Operon system (Table S1b in the Supporting Information), expression of the *DsRed* gene is also consistent with independent control in response to Amm. However, expression of the downstream *eGFP* gene within the Operon construct reveals inducer-response ranges and Hill coefficients which suggest a control mechanism that is dependent on both inducers. The results presented here for the Operon construct (Figure 2d) are significant given that, to date, no natural riboswitches have been reported to function within intercistronic regions of bacterial mRNAs. However, there are several examples in nature in which tandem riboswitches

within the 5'-UTR sequence have been shown to regulate a single gene output,^[19] which provides greater control over the dynamic range of inducer response than is possible with a single riboswitch acting alone. The two riboswitches within our Operon construct also permit a greater level of control over the dynamic range of inducer response, which allows the expansion or contraction of this range through co-administration of defined amounts of the two inducer compounds, to achieve an optimal output of the downstream gene product (Figure S3 in the Supporting Information).

To demonstrate how mutually orthogonal riboswitches can regulate the stoichiometry of two gene products, the normalized gene expression data from both the Dual Promoter and Operon systems were used to generate 2D contour plots of the relative expression landscapes (Figure 4). From these contour plots it is possible to see how appropriate inducer concentrations can be selected so that the desired coexpression ratio of two gene products can be obtained. At equivalent concentrations of Amm and 2AP, the normalized

expression ratio is 1:1 for the Dual Promoter construct (Figure 4a). Within the Operon construct this 1:1 ratio shifts towards expression of the *DsRed* gene at low concentrations (4–32 μM) of 2AP (Figure 4b). In principle, such contour plots can be used to assess the optimal coexpression ratios for any proteins of interest, for example, the coexpression of protein binding partners in a multiprotein complex, differential enzyme expression to control flux through an engineered biosynthetic pathway, or the coexpression of proteins with ancillary factors, such as chaperones, glycosyltransferases, or other posttranslational machinery.

Finally, fluorescence microscopy was used to ascertain whether the independent expression of the two reporter genes occurs within individual bacterial cells. The microscopy images of *E. coli* cells transformed with the Dual Promoter or Operon constructs (Figure S4a and S4b in the Supporting Information, respectively) show that the individual cells emit both red and green fluorescence upon co-administration of Amm and 2AP (125 μM), and that co-localization of fluorescence is clearly seen in the overlay frame (bottom right). This confirms that coexpression of DsRed and eGFP occurs within individual *E. coli* cells for both constructs, and that the expression profiles are not a result of differential expression from sub-populations of cells. Additionally, time-lapse microscopy images were acquired over a period of 8 h, after induction under the same conditions (Movie S1 in the Supporting Information).

By combining mutually-orthogonal riboswitches, a multi-component Dual Promoter coexpression system has been constructed, which can be used in combination with distinct small-molecule inducer compounds to control the differential expression of multiple genes within the same cell. The Dual Promoter system permits simultaneous and independent control over two genes with no “cross-talk” between the genetic control elements. A synthetic Operon system, in which riboswitches are deployed in both the 5'-UTR and intergenic region of a bicistronic mRNA, was shown to afford a greater level of control over the dynamic range of inducer response, wherein the dynamic range of control that is afforded by the downstream intergenic riboswitch is controlled by the upstream 5'-UTR riboswitch (Figure S3 in the Supporting Information). Furthermore, fluorescence microscopy demonstrates that coexpression occurs within the same cells. By using 2D contour plots of the dynamic expression landscapes, it is possible to select the appropriate concentrations of small-molecule inducers that are required to obtain specific coexpression ratios.

Nature has evolved natural riboswitches over billions of years that provide orthogonal control of gene expression in response to natural metabolites. However, this proof-of-principle study demonstrates that artificial constructs comprising engineered riboswitches can be used to control the expression of multiple genes of interest by using synthetic ligands, as opposed to endogenous metabolites. The mutually orthogonal riboswitches described here have induction factors (ca. 10–25) which are lower than some protein-based transcriptional expression systems. However, in this study we demonstrate that the relatively simple protein-free riboswitch regulatory elements are amenable to engineering, such that,

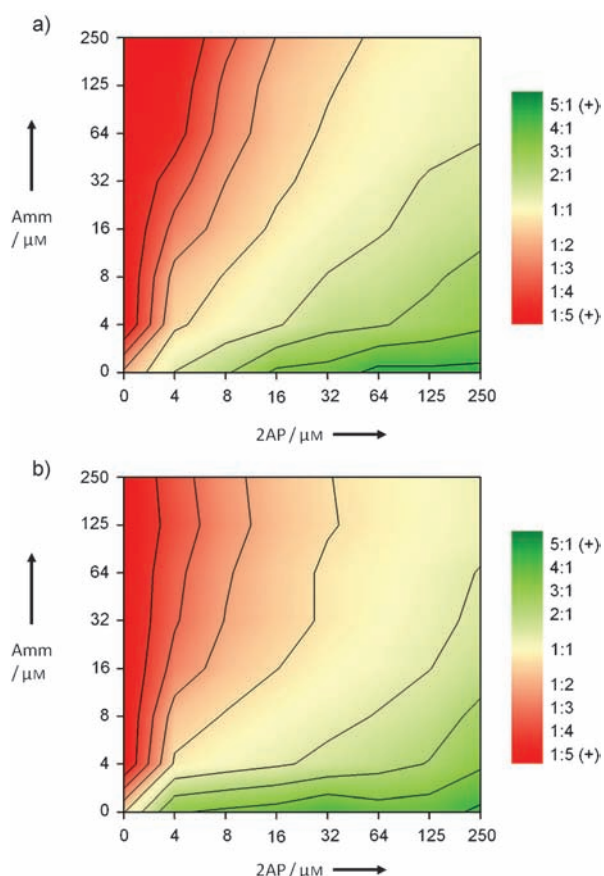


Figure 4. Contour plots of relative gene coexpression. Amm and 2AP were co-administered to cells over a range of concentrations (0–250 μM each). eGFP and DsRed fluorescence were monitored and normalized for cell density. Fluorescence values were then normalized between 0 and 1 by using the minimum and maximum expression levels determined for the Dual promoter and Operon transformants (Table S1 in the Supporting Information). Relative expression ratios (eGFP:DsRed) were calculated for the full range of Amm and 2AP concentrations tested. The results are plotted as 2D contour plots for a) the Dual Promoter system, and b) the Operon system.

with further optimization, these systems may find uses in studying gene networks, for applications of protein coexpression, and in synthetic biology.

Received: December 23, 2011

Revised: February 3, 2012

Published online: March 1, 2012

Keywords: gene expression · riboswitches · RNA · synthetic biology

- [1] S. Jana, J. K. Deb, *Appl. Microbiol. Biotechnol.* **2005**, *67*, 289–298.
- [2] R. Bertram, W. Hillen, *Microb. Biotechnol.* **2008**, *1*, 2–16.
- [3] a) Y. Ji, B. Zhang, S. F. Van Horn, P. Warren, G. Woodnutt, M. K. Burnham, M. Rosenberg, *Science* **2001**, *293*, 2266–2269; b) J. Wang et al. *Nature* **2006**, *441*, 358–361, see Supporting Information.
- [4] a) T. Ellis, X. Wang, J. J. Collins, *Nat. Biotechnol.* **2009**, *27*, 465–471; b) A. S. Khalil, J. J. Collins, *Nat. Rev. Genet.* **2010**, *11*, 367–379; c) P. E. Purnick, R. Weiss, *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 410–422.
- [5] a) J. D. Keasling, *Science* **2010**, *330*, 1355–1358; b) V. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nat. Biotechnol.* **2003**, *21*, 796–802; c) E. J. Steen, Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. Del Cardayre, J. D. Keasling, *Nature* **2010**, *463*, 559–562.
- [6] a) A. Khlebnikov, J. D. Keasling, *Biotechnol. Prog.* **2002**, *18*, 672–674; b) R. M. Morgan-Kiss, C. Wadler, J. E. Cronan, Jr., *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7373–7377.
- [7] a) J. D. Keasling, *ACS Chem. Biol.* **2008**, *3*, 64–76; b) S. K. Lee, H. H. Chou, B. F. Pfeleger, J. D. Newman, Y. Yoshikuni, J. D. Keasling, *Appl. Environ. Microbiol.* **2007**, *73*, 5711–5715.
- [8] C. Bieniossek, Y. Nie, D. Frey, N. Olieric, C. Schaffitzel, I. Collinson, C. Romier, P. Berger, T. J. Richmond, M. O. Steinmetz, I. Berger, *Nat. Methods* **2009**, *6*, 447–450.
- [9] Y. Chang, D. Mead, V. Dhodda, P. Brumm, B. G. Fox, *Protein Sci.* **2009**, *18*, 2316–2325.
- [10] B. F. Pfeleger, D. J. Pitera, C. D. Smolke, J. D. Keasling, *Nat. Biotechnol.* **2006**, *24*, 1027–1032.
- [11] N. Dixon, J. N. Duncan, T. Geerlings, M. S. Dunstan, J. E. McCarthy, D. Leys, J. Micklefield, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2830–2835.
- [12] a) M. Mandal, R. R. Breaker, *Nat. Struct. Mol. Biol.* **2004**, *11*, 29–35; b) A. Roth, R. R. Breaker, *Annu. Rev. Biochem.* **2009**, *78*, 305–334.
- [13] J. E. Graham, *Nucleic Acids Res.* **2004**, *32*, 3093–3100.
- [14] a) S. D. Gilbert, F. E. Reyes, A. L. Edwards, R. T. Batey, *Structure* **2009**, *17*, 857–868; b) A. Serganov, Y. R. Yuan, O. Pikovskaya, A. Polonskaia, L. Malinina, A. T. Phan, C. Hobartner, R. Micura, R. R. Breaker, D. J. Patel, *Chem. Biol.* **2004**, *11*, 1729–1741.
- [15] A. Deana, J. G. Belasco, *Genes Dev.* **2005**, *19*, 2526–2533.
- [16] M. H. de Smit, P. W. Verlaan, J. van Duin, C. W. Pleij, *J. Mol. Biol.* **2009**, *385*, 733–747.
- [17] a) B. M. Burmann, K. Schweimer, X. Luo, M. C. Wahl, B. L. Stitt, M. E. Gottesman, P. Rosch, *Science* **2010**, *328*, 501–504; b) V. Epshtein, D. Dutta, J. Wade, E. Nudler, *Nature* **2010**, *463*, 245–249.
- [18] J. M. Peters, R. A. Mooney, P. F. Kuan, J. L. Rowland, S. Keles, R. Landick, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15406–15411.
- [19] a) N. Sudarsan, M. C. Hammond, K. F. Block, R. Welz, J. E. Barrick, A. Roth, R. R. Breaker, *Science* **2006**, *314*, 300–304; b) R. Welz, R. R. Breaker, *RNA* **2007**, *13*, 573–582.